MPEG4에서의 텍스춰 및 모형정보를 고려한 새로운 비트율 변환기

이 종 배[†]

요 약

일반적으로 대부분의 디지털 영상처리를 위하여 현재 MPEG2[1](일부분의 경우에는 MPEG1[2])를 이용하고 있고, 이들 처리 결과가 컴퓨터에 비트스트림으로 변환되어 저장되고 있다. 그런데 전송채널은 다양한 용량을 갖고 있으며, 이에 따라 컴퓨터에 저장된 비트스트림들도 채널의 특성에 맞도록 비트율이 재조정되어야 한다. 즉 이들 비트스트림을 채널용량에 맞도록 어떻게 비트율을 바꿀 지가 이 논문에서 다루고자하는 핵심이며 이를 통상 비트율 변환이라 부른다. 실제로 지금까지 MPEG2 비트스트림에 대해서는 여러 학자들이 이들을 효율적으로 다양한 채널로 전송하기 위해서 비트율 변환을 행하는 방식들을 제안하였다. 그러나 기존의 방식들은 MPEG4 응용에는 적합하지 않다. 이는 MPEG4에서는 텍스춰 정보 뿐만이 아니라 MPEG2에는 불필요한 모형정보까지 전송해야 하기 때문이다. 따라서 본 논문에서는 텍스춰 정보와 함께 모형정보를 고려한 새로운 비트율 변환기를 제안하고자 한다.

A New Bit-rate Converter for Considering Texture and Shape Information in MPEG4

Jong-Bae Lee[†]

ABSTRACT

MPEG2(or some situations, MPEG1) is mostly used to change the raw image into the compressed bitstreams for digital image processing. Also these bitsreams are stored in computer for several applications such as multimedia area. Transmission channels have various capacities according to the application area, and the bitstreams stored in computer should be converted in order not to exceed the capacities of a transmission channel. So the problem is how to convert compressed bitsreams of a given bit-rate into compressed bitsreams of other bit-rates. Such a specific transcoding problem in this paper is referred to as bit-rate conversion. Several researches have been done on bit-rate conversion for the bitstreams compressed by MPEG2. But the existing schemes are not suitable for MPEG4 applications because it needs to transmit shape information(which is not necessary for MPEG2) as well as texture information in MPEG4 system. So we propose a new bit-rate converter which considers both texture and shape information.

1. 서 론

MPEG1이나 MPEG2 영상압축 방식은 다양한 영상 서비스를 위해 영상을 부호화하여 비트스트림 형태로 저장매체에 저장해 두고, 이를 응용에 맞도록 활용하는 방식(off-line 전송방식)이 각광을 받고 있다. MPEG2로 저장한 비트스트림은 일반적으로 고화질로 재생이 가능하도록 구성되어 있어서 데이터량이 방대하다. 따라서 채널 용량이 이를 수용하지 못할 경우 이들 데이터를 강제적으로 줄여 보내야하며, 실제로 이러한 부분에 대한 연구가 여러 학자에 의해 이루어져왔다. 그런데 향후 멀티미디어 환경 하

[†] 정회원, 명지전문대학 컴퓨터과

에서는 MPEG2 뿐만이 아니라 MPEG4의 경우도 영상 처리를 위해 널리 사용된 것이므로 MPEG4의 비트스트림으로 각종 데이터를 저장한 경우에 대해서도 비트율 변환을 위한 연구가 필요하다. MPEG4[3]는 MPEG2와는 달리 VO(video object)를 표현하기 위한 모형(shape) 정보가 필요하므로, 비트율 변환시 모형정보를 어떻게 다룰 것인지도 함께 고려하여야 한다. 이를 위해서 본 논문에서는 텍스춰 정보 뿐만이 아니라 모형 정보를 고려하여 비트율 변환을 할 수 있는 MPEG4를 위한 새로운 비트율 변환기를 제안한다.

일반적으로 부호화된 영상신호의 비트스트림을 통신채널을 통해 전송하고자 할 때, 해당 채널의 전송용량은 고정된 경우보다 가변적인 경우가 보통이다. 즉 영상을 고화질로 재생할 수 있도록 제작된 비트스트림이 다양한 전송용량을 갖는 통신채널로 전송될 경우에 더욱 그렇다. 그러므로 부호화 과정시 채널 특징을 미리 파악하여야 각 채널별로 올바른 비트스트림 전송이 행해질 수 있다. 이러한 상황에 부합하도록 본 논문에서는 MPEG4로 미리 부호화된 비트스트림을, 다양한 용량을 가진 여러 채널들로 전송하는 문제를 다루기로 하겠다. 기본적으로 비트율 변환은 낮은 지연과 저비용의 구조가 되어야하는 점에서 기존 비트율 변환기와 차이가 없지만 시스템 구현 측면에서 본 방식은 텍스춰 정보 이외에 기존 방식들[4-10]에서 다루지 않는 모형정보도 함께 고려하여 보다 효율적인 비트율 변환기를 구현하고자 한다. 제 2장에서는 먼저 다양한 기존 비트율 변환기에 대해 살펴보고 제 3장에서는 이들 기존방식들의 문제점을 개선한 새로운 방식을 제안한다. 제 4장에서는 기존 방식과 제안한 방식을 모의 실험을 통해서 비교해보고 제 5장에서 최종 결론을 도출하기로 한다.

2. 기존 비트율 변환기법들에 대한 고찰

기존의 비트율 변환기는 모두 텍스춰 정보에 대한 비트율 변환을 고려했으므로 이들 관점에서 고찰하기로 한다. 기부호화된 비트스트림을 화소 영역에서 복호화를 행하고 양자화 스텝 사이즈를 변경하여 다시 부호화를 행하는 방식이다[4-7]. 이는 복호화와 부호화 전과정을 다시 수행해야 하므로 하드웨어 구현이 어렵고, 따라서 큰 시간적 지연을 요한다. 다음 DCT 영역에서 부호화하는 방식으로 가변길이 부호화(variable length coding), 양자화(quantization) 및 역양자화(dequantization) 과정이 요구된다[8]. 이 방식은 화소영역에서 부호화하는 방식에 비해서는 하드웨어 복잡도는 줄어들면서 시간적 지연도 줄어든다. 마지막으로 양자화된 DCT 영역에서 비트율 변환을 행하는 방식으로 가변부호화 과정만이 필요하며, 하드웨어 복잡도도 작고 시간적 지연도 가장 작다[9].

그림 1은 전형적인 비트율 변환기의 일례를 나타낸 것으로 위치 (a)가 양자화된 DCT 영역에서 비트율 변환이 이루어지는 기준점이 되며, 위치 (b), 위치 (c)는 각각 DCT 영역과 양자화 DCT 영역에서의 비트율 변환의 기준점이 된다. 결국 기존의 여러 방식들은 기본적으로 이 세 가지 형태중의 하나를 취한다.

Keesman등[4]은 그림 2와 같이 화소 영역에서 비트율 변환기를 구현하였는데, 기존의 움직임 벡터와 기존 decision 값을 부호시에 그냥 이용함으로써 어느 정도 하드웨어 구현비용을 줄일 수 있었다. 그림 3에서와 같이 DCT 영역에서 재양자화를 행하는 방식의 일례로서 Nakajima[8]는 양자화 스텝 사이즈를

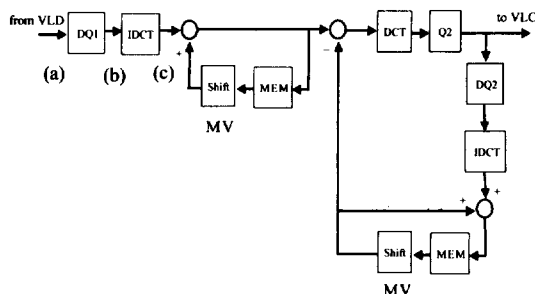



그림 1. 비트율 변환기의 블록도

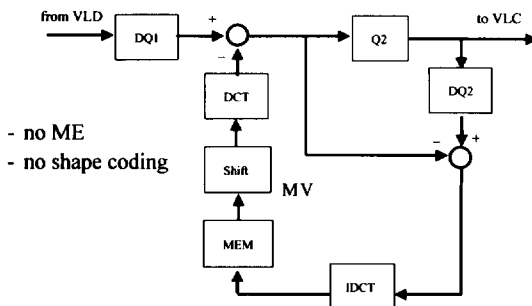



그림 2. 화소영역에서의 비트율 변환기의 구현

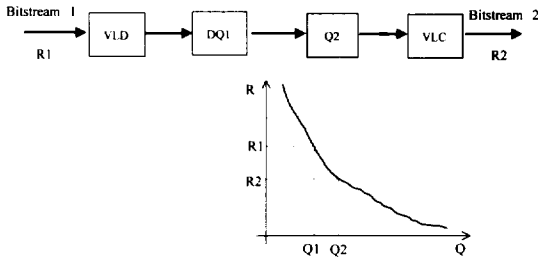



그림 3. DCT 영역에서의 비트율 변환기 구현

키우므로써 전송 비트율을 줄이는 방식을 제안하였다. 이 경우는 움직임 벡터 예측이 필요하지 않고 DCT도 필요하지 않다. 즉 재양자화 시의 스텝사이즈 Q2를 원래의 비트스트림으로부터 복원한 신호 Q1에 비해서 크게 함으로써 비트율을 줄일 수 있다. 양자화 DCT 영역에서 비트율 변환기를 구현하는 방식의 일례로서 H. sun 등[9]은 그림 4에서와 같이 AC 계수를 조절함으로써 비트율을 가변하는 방식을 제안하였다. 여기서는 움직임 예측, DCT 뿐만이 아니라 양자화도 불필요하게 되어 가장 구현이 간단하고 전송지연 또한 작다. 이들 이외에도 이들 세 방식을 적절히 조절해서 사용하는 방식[10]도 제안되고 있다.

표 1은 기존의 여러 방식들에 대해서 하드웨어 구현측면과 PSNR을 비교한 결과를 보여주고 있다. 여기서 보면 일반적으로 하드웨어 복잡도가 증가함에 따라 복원된 영상의 화질이 개선됨을 알 수 있다. 즉 응용분야에 따라 성능과 하드웨어 복잡도 간의 적절한 조화(compromise)가 필요하다.

비트율 변환기를 다시한번 간단히 표현하면 그림

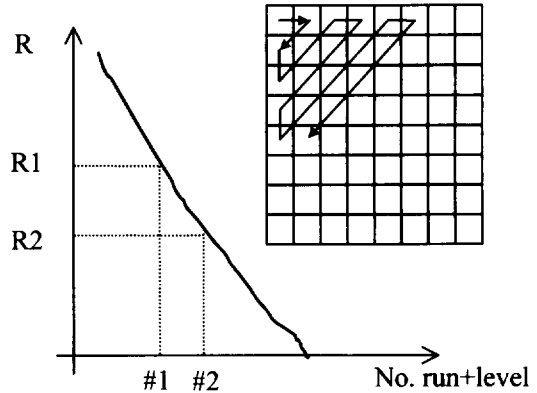



그림 4. 양자화 DCT 영역에서의 비트율 변환기의 구현

5와 같다. 즉 고화질로 부호화되어 저장된 비트스트림 R1을 비트율 변환기에서 R2로 변환하여 전송하게 된다. 즉 비트율 변환기 구성에서 중요하게 고려되어야 할 사항은 망의 폭주를 방지할 수 있도록 전송비트율을 적절히 조절하는 문제와 비용절감을 위한 하드웨어 단순화, 마지막으로 사용자의 요구에 맞도록 정보량을 자유롭게 조절할 수 있도록 하는 등의

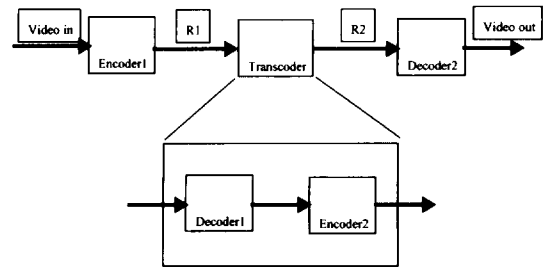



그림 5. 비트율 변환기의 구성

표 1. 기존 방식들의 성능 비교

	Pixel	Pixel	DCT	Q+DCT
Hardware Complexity	no ME	no ME no Decision	no ME/MC no Encoding loop no decoding loop no frame memory no DCT/IDCT simple rate control	no ME/MC no encoding loop no decoding loop no frame memory no DCT/IDCT simple rate control no Q/DQ no VLC
PSNR (Flower)	29.32	29.19	27.41	20.89
PSNR (bicycle)	27.11	27.02	25.04	20.18

문제이다. 앞에서 언급했던 것처럼 기존의 방식들은 주로 MPEG1이나 MPEG2의 응용을 위해 제안된 방식이므로 텍스춰 정보만을 고려하였다. 실제로 이들 기법들은 MPEG4의 특별한 경우인 한 프레임에 물체가 하나 있는 경우에는 제안한 방식과 큰 차이없이 응용될 수 있다. 그런데 일반적으로 한 프레임에는 여러 물체가 존재하므로 경계정보를 함께 고려하는 것이 보다 타당하다.

3. 제안한 방식

본 논문에서는 물체의 모형정보(shape information)를 고려한, MPEG4에 맞는 새로운 비트율 변환기를 제안한다. 즉 텍스춰 정보와 모형정보를 모두 고려하고 VO(video object) 단위의 제어를 행하며, 다양한 멀티미디어 데이터의 효율적 응용이 가능하도록 한다.

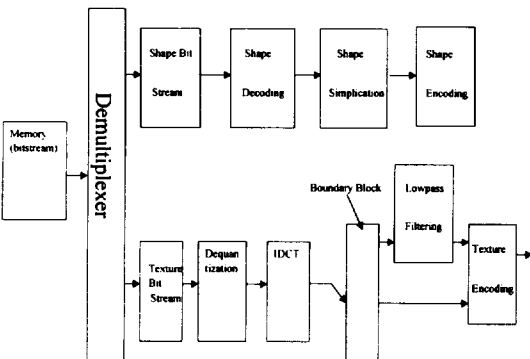
그림 6은 제안한 방식의 간단한 블록도를 나타내고 있다. MPEG4로 부호화된 비트스트림은 메모리에 off-line 형태로 저장되어 있게 된다. 이들 데이터는 모형정보, 움직임 정보, 텍스춰 정보로 구성되어 있는데 움직임 정보는 그 중요성에 비추어 전송비트를 감축하는 것이 실익이 없으므로 비트율 변환시에도 그대로 전송된다. 즉 움직임 정보를 감축하여 전송하면 각 VO 평면 사이의 텍스춰 정보가 증가하여(각 VOP 간 움직임 정보의 부정확성으로 인하여 움직임 보상 오차가 증가하게 됨) 급격한 화질 저하를 초래하므로, 대부분의 경우에 전송시 비트 감축없이 그대로 보낸다. 본 방식에서도 움직임 정보는 디멀티플렉서에서 분리되어 비트 감축없이 그대로 전송되므로 그림 6에서는 생략하였다. 다음에 모형 비트스트림 정보는 모형 복호화부(shape decoding)에서 복호화 과정을 거쳐 원래의 모형이 복원된다. 복원된 모형정보는 모형개략화부(shape simplification)에서 모형단순화 과정을 거쳐 블록단위(블럭크기는 응용분야에 따라 적절히 조절하게됨)로 재부호화함으로써 모형 발생비트를 줄이게 된다. 텍스춰 정보는 화소 단위의 비트율 감축방식을 사용하였는데, 먼저 디멀티플렉서에서 분리된 텍스춰 비트스트림을 입력으로하여 역양자화부(Dequantization)에서 역양자화를 행하고 IDCT부에서 IDCT를 행하여 텍스춰를 복원한다. 복원된 테스춰 데이터는 모형정보가 존재하는 블럭(boundary block)과 내부블럭으로 구분하여, 내부블럭의 경우에는 양자화 스텝 사이즈를 변경하여 부호화하고 경계블럭의 경우에는 물체의 경계가 찌그러지는 현상을 방지하기 위하여 저역통과 필터(low pass filter)를 통과시킨 후에 텍스춰 부호화(texture encoding)부에서 텍스춰 부호화 과정을 행하게 된다.

4. 실험결과 및 고찰

그림 7은 MPEG-4 표준영상인 Akiyo 영상에서의 배경에 해당하는 obj0에 대하여 obj0_shape_pixel (MPEG-4 적용)과 obj0_shape_pixel(4×4)(shape정보 표현시 4×4 블록까지 세분화한 경우) 실험결과를 제시하였고, 또 얼굴영상에 해당하는 obj1_shape_pixel과 obj1_shape_pixel(4x4)에 대하여 마찬가지로 실험결과를 나타낸 것이다. 실험결과를 통

그림 6. 제안한 방식의 블록도

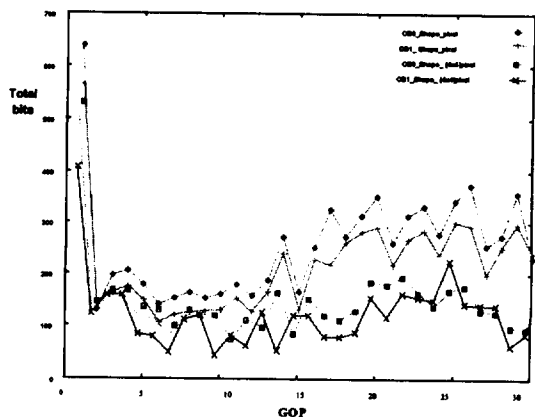



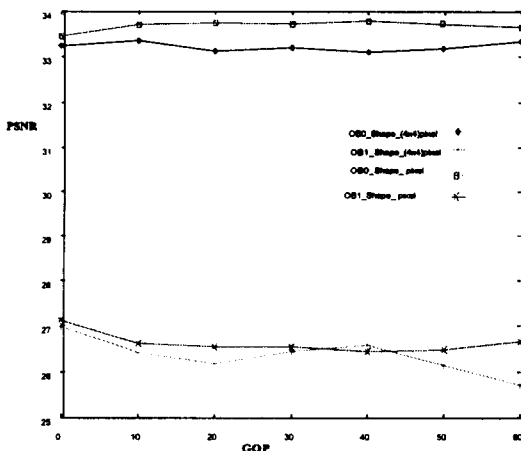
그림 7. 모형정보의 비트 발생량 비교

해서 알 수 있듯이 object의 경계(boundary)를 픽셀 단위까지 처리하여 표시한 경우에 비해서 4×4 블록까지 세분화한 경우가 데이터 발생량이 약 30~40% 감소하는 것을 확인할 수 있었다.

그림 8은 PSNR 측면에서의 모형 단순화 과정을 통한 모형정보의 감축을 나타내고 있다. 즉 Akiyo 연속영상에 대하여 각 object 별(OBJ-0, OBJ-1)로 픽셀단위와 4×4 블록 단위까지 PSNR 측면에서 모의실험을 행한 결과이다. 결과를 살펴보면 4×4 블록 단위까지 모형정보를 표시한 경우가 픽셀단위까지 모형정보를 모두 표시한 것에 비해서 약 0.2~0.3 dB 정도 PSNR이 감소하였음을 알 수 있다. 이는 전송비트의 감소를 고려하면 충분히 허용할 수 있는 화질의 열화로 판단되며, 적어도 화질의 객관적 화질 측면에서는 열화가 큰 문제가 되지 않는 것을 알 수 있었다. 이를 30프레임, 60프레임 단위로 OBJ-0와 OBJ-1에 대해서 모형정보의 전송에 필요한 비트를 MPEG4와 제안한 방식에 대해서 구체적으로 표시하면 아래와 같다.

a. 프레임(Akiyo 연속영상)에 대한 모형정보 발생 비트

화소단위의 모형분할의 경우
- OBJ-0 : 7630비트(Y에 대한 PSNR : 33.67[dB])
- OBJ-1 : 6487비트(Y에 대한 PSNR : 26.64[dB])

(4×4) 블록 단위의 모형분할의 경우
- OBJ-0 : 4849비트(Y에 대한 PSNR : 33.19[dB])
- OBJ-1 : 4248비트(Y에 대한 PSNR : 26.43[dB])

b. 60 프레임(Akiyo 연속영상)에 대한 모형정보 발생비트

화소단위의 모형 분할의 경우
- OBJ-0 : 15553비트(Y에 대한 PSNR : 33.70[dB])
- OBJ-1 : 13687비트(Y에 대한 PSNR : 26.56[dB])

(4×4) 블록 단위의 모형 분할의 경우
- OBJ-0 : 8566비트(Y에 대한 PSNR : 33.18[dB])
- OBJ-1 : 7376비트(Y에 대한 PSNR : 26.34[dB])

이 결과로부터 모형을 블록단위로 개략화한 제안한 방식이 화소 단위의 모형분할 방식에 비해서 비트 감축이 획기적으로 이루어짐을 알 수 있다. 특히 30 프레임의 경우보다 60 프레임의 경우가 비트 감축이 더 많이 이루어짐을 알 수 있는데, 이는 시스템이 초기화 과정을 거쳐 보다 안정된 상태에서 제안한 방식이 더 큰 효과를 나타내기 때문이다.

그림 9는 QP의 변화에 따른 텍스춰/모형 정보의 상대적 비율을 표시하고 있다. 실험결과에서 QP가 클수록(데이터 감축을 많이 할수록) 모형정보가 전송에서 차지하는 비중이 커진다. 따라서 클라이언트 측에서는 상대적으로 서버 측보다는 데이터 압축을 많이 하여야 한다. 즉 QP 값을 상대적으로 크게 하여 데이터를 감축하게 되는데, 이 경우 모형 전송정보가 차지하는 비중이 커지므로 전송채널의 용량이 작을수록 모형 정보의 감축 필요성이 증대함을 알 수 있었다.

그림 8. PSNR에서의 성능비교

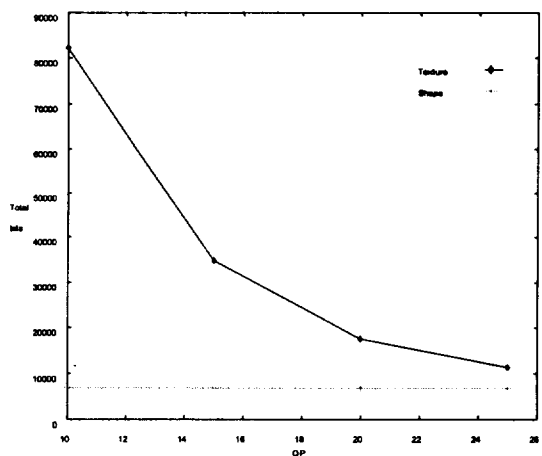



그림 9. QP의 변화에 대한 Texture/모형정보의 상대적 비율

5. 결 론

본 논문에서는 모형정보의 감축이 필요함을 입증해 보이기 위해서 위와 같이 모형 정보를 MPEG-4처럼 픽셀단위까지 표현하지 않고 4×4 블록까지 표현하였다. 결과를 보면 모형정보가 기존 기법에 비해서 대폭 감소함을 알 수 있으며, 따라서 텍스춰 뿐만이 아니라 모형정보까지 고려하여 트랜스코딩을 행할 필요가 있음을 알 수 있다. 따라서 본 연구의 결과로서, MPEG-4에 맞는 비트율 변환기는 텍스춰 정보와 모형정보를 고려하고 또 VO 단위의 제어를 행하며 다양한 멀티미디어 데이터의 효율적 접근을 지향하는 방향으로 개발되어야 함을 제시해주고 있다. 실제로 이러한 비트율 변환은 멀티미디어 환경 하에서는 갈수록 그 중요도가 더해가고 있다. 따라서 이러한 연구는 실제 정보화 사회를 구현하는데 중요한 역할을 하리라 판단된다. 제안한 방식을 부호화 측면에서 살펴보면 모형정보의 개략화를 행하기 때문에 모형정보 발생비트 감축과 더불어 모형정보의 추출도 기존의 방식에 비해서 쉽게 할 수 있는 장점이 있다. 또 CAE(Context-based Arithmetic Coding) 구현시에도 화소단위까지 모형정보를 추출하는 것에 비해서 계산량을 줄일 수 있다.

향후 해결해야할 과제로서는 블록효과(Blocking effect)를 줄이는 문제, 경계부분의 저역 필터링을 효과적으로 행하는 방안, 그리고 재생영상에 대한 후처리를 통해 재생영상의 화질을 개선하는 문제 등을 들 수 있겠다.

참 고 문 헌

[1] MPEG Video Editorial Group, "Generic coding of moving pictures and associated audio : ISO/IEC 13818-2 Draft International Standard," *ISO/IEC JTC1/SC29/WG11, MPEG94/No. 702,* Mar. 1990.

[2] Simulation Model Editorial Group, "MPEG Video Simulation Model Three(SM3)," *ISO/IEC JTC1/SC29/WG11, MPEG90/No.041* ,July 1990.

[3] AOE Group, "Coding of moving pictures and associated audio : MPEG-4 Committee Draft," *ISO/IEC JTC1/SC29/WG11, N2202,* Mar. 1998.

[4] G. Keesman, R. Hellinghuizen, and Fokke Hoeksema, and Geert Heideman, "Transcoding of MPEG bitstreams," *Signal Processing: Image Communication 8,* pp. 481-500, 1996.

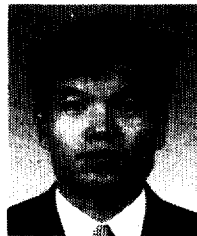
[5] Pedro A. A. Assuncao, "Post-processing of MPEG-2 video for transmission at lower bit rates," Proc. ICASSP, 1996.

[6] D. G. Morrison, M. E. Nilsson, and M. Ghanbari, "Reduction of bit-rate of compressed video while in its coded form," in *Sixth International Workshop on Packet Video,* pp. 392-406, Sep. 1994.

[7] G. Morrison, "Video Transcoders with Low Delay," *IEICE. Trans. Communi.,* Vol. E80-B. No. 6, pp. 963-969, Jun. 1997.

[8] Yasuyuki Nakajima, Hironao Hori, and Tamotsu Kanoh, "Rate conversion of MPEG coded video by requantization process," *Proc. ICIP,* pp. 408-411, 1995.

[9] H. Sun, W. Kwok, and J. W. Zdepski, "Architectures for MPEG compressed bitstream scaling," *IEEE Trans. on Circuits and Systems for Video Technology,* Vol. 6, No. 2, pp. 191-199, April 1996.

[10] 서광덕, 김재균등, "재양자화 특성을 이용한 비트율 변환기의 전송률 제어 기법," *방송공학회 학술대회* 1997, pp. 181-186, 1997.

이 종 배

1985년 2월 서울대학교 제어계측학과 졸업(학사)
1987년 2월 KAIST 전기 및 전자공학과 졸업(석사)
1996년 2월 KAIST 전기 및 전자공학과 졸업(박사)
1985년 3월~1996년 9월 LG전자 연구원
1996년 10월~1997년 8월 특허청 영상기기과 심사관
1997년 9월~현재 명지전문대학 컴퓨터과 조교수